Individual Recognition Based On Transfer Learning For Wireless Network Device

Zhibing Yu, Ruilun Zong and Shuangqiu Li

October 14, 2019

Individual Recognition Based On Transfer Learning For Wireless Network Devices

Zhibin YU
Southwest Jiaotong University
Xi'an Road, Pidu District,Chengdu
Sichuan, China
+86-18224075056
zbyu@swjtu.edu.cn

Ruilun ZONG
Southwest Jiaotong University
Xi'an Road, Pidu District,Chengdu
Sichuan, China
+86-17358653392
zongruilun@foxmail.com

Shuangqiu LI
Southwest Jiaotong University
Xi'an Road, Pidu District,Chengdu
Sichuan, China
+86-17361061979
lsq19950222@163.com

ABSTRACT
Due to the complexity of the electromagnetic environment and the non-cooperative characteristics of the target in the process of individual recognition of wireless network devices, the amount of data acquired is small, and the data required for training the model cannot be enough. For this small samples situation, a new method of individual recognition of wireless network devices based on transfer learning is proposed. We combine a convenient signal processing method with transfer learning in this paper. The acquired signal is divided into source domain and target domain. We extract the feature of transient signal and get the weight of feature.Then we transfer the knowledge of the source domain to the target domain. The training model is reconstructed by adjusting the weight of the source domain samples, which solves the problem of insufficient training samples in the small samples situation. The results show that in the case where the number of target domain samples is less than 30, the recognition rate has improved obviously.It enriches the application of transfer learning in the field of individual recognition.

CCS Concepts
Networks → Network protocols → Wireless frame signal

→ identify wireless network devices

Keywords
Individual recognition; Small samples; Transfer learning

1. INTRODUCTION
Individual recognition of wireless network devices is a research hotspot in recent years. It is of great significance in the fields of radio security communication, military communication confrontation and civil radio monitoring.

At present, researchers at home and abroad mainly recognize individual wireless network devices from the subtle features of transient signals (referred to as transient characteristics) and the subtle features of steady-state signals (referred to as steady-state features). In 2011, transient signal fingerprints were applied to recognize IEEE802.11b wireless network devices, achieved 90% correct recognition rate[1]. At the same time, the method of extracting radio frequency fingerprinting by bispectrum analysis and phase space reconstruction is proposed, and the clock frequency offset fingerprint is used for synchronous digital system network device recognition[2]. In 2014, steady-state signal fingerprints were applied in relay authentication and recognition[3]. In 2015, Knox et al. used steady-state signal phase information as the RF fingerprint to recognize the same manufacturer devices, with an average classification accuracy of 81.9%[4].

Actually, not only do we need to face a variety of electromagnetic interferences, but also the targets that need to be recognized are often non-cooperative, the amount of data acquired is small, and the individual recognition of wireless network devices becomes more difficult. Currently,there are few studies in this area.

In view of the small amount of acquired data and the inability to build the model, transfer learning as a machine learning method is very suitable for solving this problem, but there are few reports that the application of transfer learning in the individual recognition of wireless network devices. In this paper, we not only use a convenient signal processing method but also transfer the knowledge of the source domain to the target domain. We can get a better result through experimental verification.

2. RESEARCH FRAMEWORK
Through the analysis of research results at home and abroad in recent years, in general, we find that the transient signal is simpler than the steady-state signal, the hardware requirements are lower and the expected effect is better. Therefore, this paper chooses to study the transient signal.

Transient signals are generated from the instantaneous impact of hardware components in wireless communication, so each wireless communication device can generate transient signals. Due to the nuances of the component processing technology and the relative position of the components during the loading process, the transient signals of each wireless communication device are different. In addition to the difference in working time and working environment, this difference will be more obvious.So it can be used to reflect the characteristic of the wireless network devices.

The transient signal has a short duration and changes drastically. Therefore, the detection of the starting point of the transient signal is very important. At present, the detection methods include phase detection method[5] and short-time energy detection method[6]. The SNR is large in the experimental environment of this paper, so short-time energy detection is chosen.

The transient signal obtained for the first time exists in the form of a two-dimensional image. Feature extraction and quadratic feature extraction can be performed to obtain a one-dimensional feature vector, which facilitates subsequent works.

All of the above is the basis for individual recognition of wireless network devices in the case of small samples. Next, the transfer learning method will be used to transfer the knowledge of the source domain to the target domain for improving the recognition rate under the small samples.

3. TRANSIENT SIGNAL PROCESSING

This section describes the signal processing method

3.1. Short-time energy method

The short-time energy detection method is a common starting point detection method that is very suitable for the initial stage energy mutation. The short-time energy detection method is based on the amplitude of the signal. The energy curve can reflect the amplitude variation of the signal well. By detecting the abrupt point of the signal energy curve, the detection of the starting point of the signal can be performed.

The short-time energy *E(n)* can be expressed as:

$$E(n) = \sum_{i=n-N+1}^{n} [x(i)w(n-1)]^2 \tag{1}$$

Where *w(n)* represents a window function, generally, a *Rectangular window* or a *Hamming window* will be used. In this paper, a *Rectangular window* function is selected.

The accuracy of the start of the signal is closely related to the width *w* of the window function, and the window width *w* determines the time at which the signal energy is integrated. Generally speaking, we do not have a unified standard for the definition of window width. In this paper, we have determined the window width *w*=100 and the value of frame shift is 5.The short-time energy curve of the available signal is calculated by formula (1). The energy is weak and remains basically unchanged in the begining. After the start of the signal, the short-time energy value begins to fluctuate. Combined with the threshold detection method[7], the threshold value τ is set. If the current point and followed 5 points by a small pitch are greater than τ, this point is considered as the starting point. The figure 1 below shows the signal and energy changes.The signal captured from the real wireless network (IEEE 802.11).

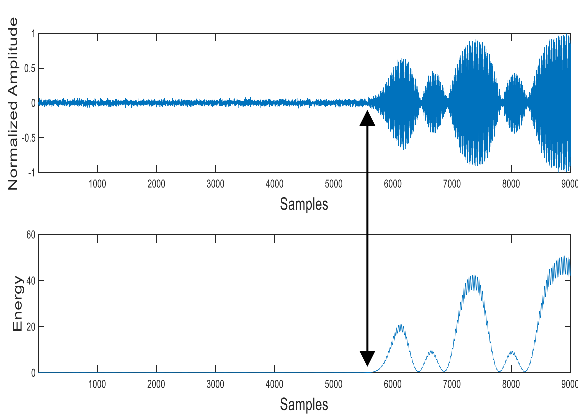

Fig 1 Signal and energy changes

3.2. Signal processing based on Hilbert and least squares fitting

The *Hilbert* transform of the real signal *x(t)* is defined as:

$$\hat{x} = x(t) * \frac{1}{\pi t} = \frac{1}{\pi} \iint_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \tag{2}$$

Then the analytical signal *g(t)* of *x(t)* can be expressed as:

$$g(t) = x(t) + j\hat{x}(t) \tag{3}$$

The amplitude of the real signal *x(t)* is equal to the amplitude of the analytical signal *g(t)*, so the envelope curve of the original real signal *x(t)* can be obtained, expressed as:

$$A(t) = |g(t)| = \sqrt{x^2(t) + \hat{x}^2(t)} \tag{4}$$

It can be seen from figure 2 that the amplitude envelope obtained by *Hilbert* transforming contains a wealth of information.

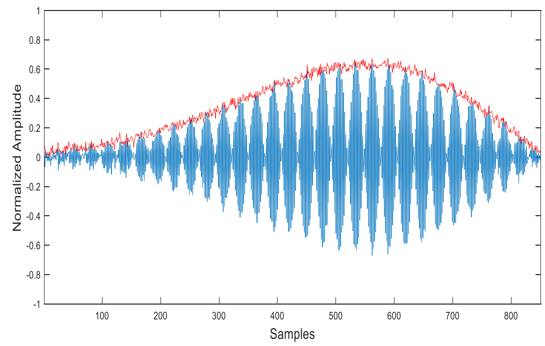

Fig 2 Signal and envelope

However, the amplitude envelope above has many burrs, and there are partial mutations, so optimization is needed. The least squares polynomial fitting method will be used.

The least squares polynomial fitting method is to find the function curve *y=P(x)* to minimize the sum of squared deviations from the actual data points *(xi, yi)*.

The fitting function *P(x)* can be expressed as

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n \tag{5}$$

Where n is the fitting order and a_i（i=1,2，…n）is the fitting function coefficient.

Then we can get:

$$\frac{\partial l}{\partial a_j} = 2\sum_{i=0}^{N} \left(\sum_{n=0}^{m} a_0 x_i^k - y_i\right) x_i^j = 0 \tag{6}$$

And

$$P(x) = \begin{cases} a_0 N + a_1 \sum x_i + \cdots + a_m \sum x_i^m = \sum y_i \\ a_0 \sum x_i + a_1 \sum x_i^2 + \cdots + a_m \sum x_i^{m+1} = \sum x_i y_i \\ \vdots \\ a_0 \sum x_i^m + a_1 \sum x_i^{m+1} + \cdots + a_m \sum x_i^{2m} = \sum x_i^m y_i \end{cases} \tag{7}$$

Solving the equation.Figure 3 shows the fitted results.

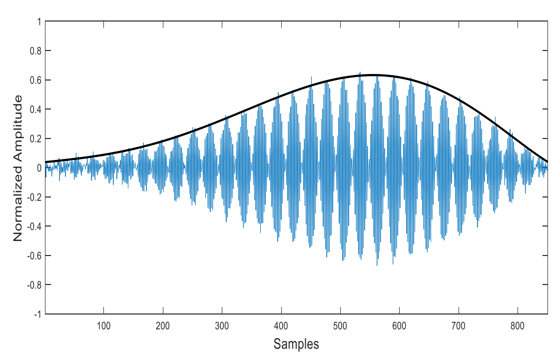

Fig 3 Signal and fitted envolope

As the digital communication system adds the preamble sequence before the data segment[8], different wireless network devices not only have different transient parts, but also have different preambles. The following figure 4 is a comparison of beacon frame headers sent by different wireless network devices collected by the network (IEEE 802.11).

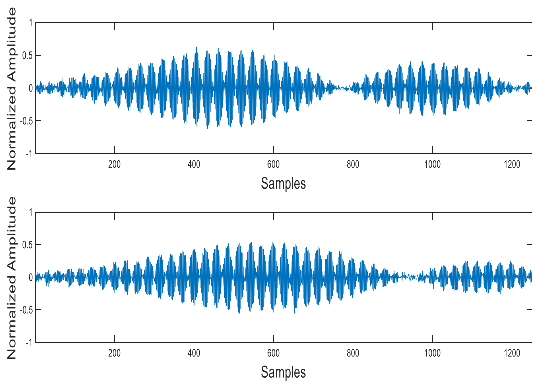

Fig 4 Comparison of frame headers sent by different wireless network devices

4. TRANSIENT ENVELOPE FEATURE EXTRACTION AND ENTROPY WEIGHTING METHOD

4.1. Transient envelope feature extraction

Individual recognition of the wireless network devices dose not need so much information, so we need to perform transient envelope feature extraction.

Here are several methods[9]:

(1) Area of the transient curve

(2) Mean of transient curve

(3) The skewness of the transient curve

(4) The kurtosis of the transient curve

(5) Variance of transient curve

(6) Maximum slope of the transient curve

(7) Box dimension of transient signals

(8) Information entropy of transient signals

4.2. Entropy weighting method to determine feature weight

The extracted features are different in importance for classification. If the importance of the features of the extraction is equal blindly, the effect of the features that best reflect the individual differences will be weakened, thereby reducing the training effect of the classifier. It will affect the recognition accuracy. So we need to determine the weight of the various features. The following is the entropy weighting method[10] process:

Entropy weighting

Step1: Measure the index of the matrix $X = (x)_{m*n}$, calculate the parameter weight of the i-th value of the j-th parameter $P_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}}$, normalize the matrix, and obtain the standard matrix $(P_{ij})_{m*n}$;

Step2: Calculate the entropy value of the j-th parameter $E_j = -\frac{1}{\ln m}\sum_{i=1}^{m} P_{ij} \ln(P_{ij})$;

Step3: Calculate the difference coefficient of the j-th parameter $d_j = \frac{1}{E_j}$;

Step4: Calculate the weight of the j-th parameter $v_j = \frac{d_j}{\sum_{j=1}^{n} d_j} (j = 1,2,\dots,n)$.

5. APPLICATION OF TRANSFER LEARNING IN THE CASE OF SMALL SAMPLES

The domain of transfer learning represents the subject of learning, expressing the characteristics of the data and the distribution of features. The domain is divided into a target domain and a source domain. The target domain refers to the domain to be learned, and the source domain refers to the domain of existing knowledge. Transfer learning is to apply the knowledge learned in the source domain to the target domain to complete the target task. Transfer learning is widely used in various fields. Transfer learning can be divided into many types, including transfer learning based on instance,transfer learning based on feature, transfer learning based on shared parameters, etc.

Due to the electromagnetic interferences, target non-cooperation, etc. Engineers often get very few signals. It will create a situation. There is less signal to provide training. The training model is inaccurate and the recognition accuracy is low. So we can consider transfering the knowledge of the source domain to improve the training model in order to achieve higher recognition accuracy.The target task schematic diagram is as follows：

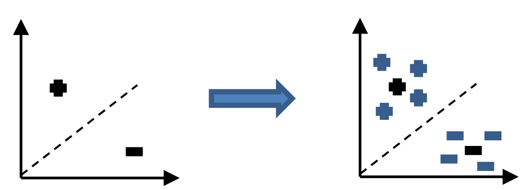

Fig 5 Target task schematic diagram

The TrAdaBoost [11] algorithm is the first instance-based transfer algorithm. The algorithm assumes that the source domain and the target domain have the same feature space. The domain difference is only the difference of data distribution. The TrAdaBoost algorithm uses AdaBoost to update the target domain weights and uses the Hedge(β) algorithm to update the source domain weights. In the iterative process, the weight of the wrong sample is reduced, and the influence of the sample on the training of the next classifier is weakened.So the knowledge can be transferred by this way.

In this paper, two wireless network devices (TpLink) with the same production date are used. The oscilloscope with a sampling rate of up to 40 GS/S is used as experimental equipment. Collect 40 sets of data as target domain samples, the source domain samples are 200 sets. We use the TrAdaBoost algorithm and the KNN classifier (k is 5).The algorithm was set to 100 iterations, and 500 Monte Carlo simulations were performed with a negative transfer of -9.65%.

For the reasons for the negative transfer, the analysis is as follows:

a. There are a few samples in the target domain in this paper, and the feature dimension is very low.The number of samples in the source domain is large and requires multiple iterations. However, after multiple iterations, the weight of the target domain has been excessively distorted, which affects the training of the classifier.

b. In the worst case, the convergence speed of the iteration is $\sqrt{ln\, n/N}$, in general, the convergence speed is $ln\, n/N$, and the convergence speed is relatively slow;

c.TrAdaboost pays too much attention to the samples that can not be recognized. The existence of these samples affects the number of iterations and affects the training.

It can be seen that this algorithm does not apply to the recognition of individual wireless network devices in the case of small samples. A new algorithm is needed to solve the problem in this paper. Now, the target domain in this algorithm is the newly collected data under IEEE802.11 network. The source domain is the previously collected data. The process is as follows:

Recognition algorithm in the case of small samples

Step1: Initialize the sample weight of the source domain (with labels) $p^t = w^t/(\sum_{i=1}^{n+m} w_i^t)$;

Step2: Obtain the training set X_{target} from the target domain and train to obtain the weak classifier h_t;

Step3: Use the source domain as a verification set and input it to the weak classifier with the purpose of obtaining sample sets X_{wrong} and X_{right};

Step4: Compare the similarity between the source domain and the target domain, set the coefficient β_0, and X_{wrong} and X_{right} update weight $w_i^{t+1} = w_i^t \pm \beta_0$, i=1, 2...m;

Step 5: Set the minimum threshold, construct a new trusted auxiliary sample X_{new} from the source domain. Merge the set $X_{new} \cup X_{target}$ and construct a new classifier.Then,adjust the minimum threshold repeatedly and optimize the training model;

Step6: Get the final classifier h_f

6. EXPERIMENT AND ANALYSIS

6.1. Experiment of individual recognition

In this paper, 5 wireless network devices (TpLink) with the different production date are used as experimental objects. The oscilloscope with a sampling rate of up to 40 GS/S is used as experimental equipment. We collected 500 sets of data and processed them. The comparison figrue is below.

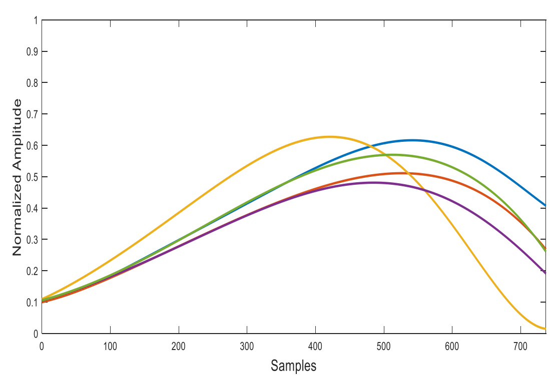

Fig 6 Comparison of 5 devices

The extracted feature from transient envelope and calculation of feature weights are shown in Table I:

TABLE I Calculation Result

Area	**Mean**	**Skewness**	**Kurtosis**
0.1226	0.1226	0.0109	0.4333
Variance	**Maximum Slope**	**Box Dimension**	**Information Entropy**
0.1038	0.0574	0.0453	0.1041

According to the weight size, the first three features skewness, the maximum slope and the information entropy are selected to establish a three-dimensional figure, as shown in figure 7.

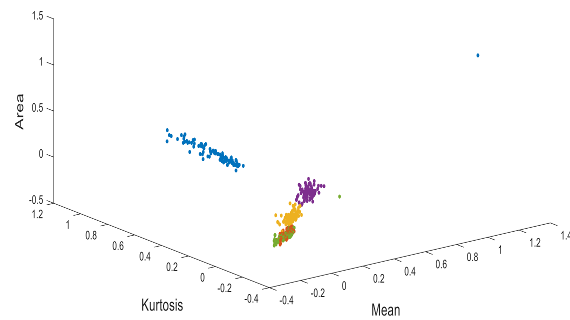

Fig 7 Distribution figure

The KNN classifier(k=5) is used for recognition, and the results are shown in Table II:

TABLE II Classification Result

Experimental Program	Recognition Rate
20% train，80% test	86.25%
30% train，70% test	88.86%
40% train，60% test	90.67%
50% train，50% test	90.80%

6.2. Experimental verification in the case of small samples

In order to reduce the time complexity of the algorithm under the small samples situation and verify the validity under the broader conditions, this paper uses four wireless network devices (TpLink) with the same production date as the experimental object, and collects 40 sets of data.Source domain is 360 sets. The KNN classifier (k is 5) is used for recognizing.We adjust the weight update of the misclassified samples X_{wrong} in the source domain by pt→0, the number of iterations is 1 time, and 500 Monte Carlo simulation experiments are performed. The experimental results are shown in the following table III.

TABLE III Classification Result

Number Of Samples	Before Transfer Learning	After Transfer Learning	Promotion
10	59.28%	72.48%	13.20%

15	69.89%	79.30%	9.41%
20	75.43%	81.71%	6.28%
25	76.49%	81.85%	5.36%
30	80.46%	84.53%	4.07%
35	83.63%	86.38%	2.75%

7. CONCLUSION

The method of individual recognition in the case of small samples in this paper enriches the application of transfer learning in the individual recognition.Under the conditions that in 10-25 sample interval,the effect is good. Under the sample interval of 30 and above, the increase is limited. Because the number of samples is enough to obtain a better training model. In summary, this method has relatively practical value.

8. ACKNOWLEDGMENT

This work is supported by the Equipment Development Department Funds grant 61403120304<JZX7Y20190252032701>) and the Fundamental Research Funds for the Central Universities grant (2019XJ04).

9. REFERENCES

[1] Zhiyuan Shi, Man Liu, Lianfen Huang, Transient-based identification of 802.11b wireless device, Wireless Communications And Signal Processing, 2011:1-5.

[2] Yuan Y, Huang Z, Wang F, Radio specific emitter identification based on nonlinear characteristics of signal, In: IEEE International Black Sea Conference on Communications and Networking, 2015:77-81.

[3] Peng H, Xianbin W, Behnad A, Relay authentication by exploiting I/Q imbalance in amplify and forward system, In: IEEE Global Communications Conference, 2014:613-618.

[4] Knox D A, Kunz T, Wireless fingerprints inside a wireless sensor network, ACM Transactions on Sensor Networks, 2015, 11(2):1-30.

[5] Hall J, barbeau M., Kranakis E, Detection of transient in radio frequency fingerprinting using signal phase . Wireless and Optical Communications. 2003.

[6] Zhao J H, Yang H. Acoustic emission source locating method based on shorttime energy. Journal of Vibration and Shoc, 2013,32(23):110-114.

[7] Wu Y C, Chen M M. Adaptive Threshold Detection Scheme in Gliding Window, 2008(10):2770-2773.

[8] Kenedy I O, Scanlon P, Mullany F J. Radio transmitter fingerprinting: A steady state frequency domain approach, In: 68th Vehicular Technology Conference, 2008:1-5.

[9] Saeed Ur Rehman, Kevin Sowerby, Colin Coghil. RF Fingerprint Extraction from the Energy Envelope of an Instantaneous Transient Signal. Australian Communications Theory Workshop (AusCTW), 2012.

[10] Li Nan. Emitter Recognition Based on Adaptive Weight Attribute Measure and Information Fusio, Journal of Projectiles,Rockets,Missiles and Guidance. http://kns.cnki.net/kcms/detail/61.1234.TJ.20190111.1032.002.html

[11] Dai W, Yang Q, Xue G R, et al. Boosting for transfer learning. Proceedings of the 24th international conference on Machine learning. ACM, 2007:193-200